**Chat 2**

**Notes 1**

**Questions?**

**6. Hypotheses**

Note table of hypotheses on page 2 of Notes 1.

Also note template and examples in supplemental document for Notes 1 videos.

**Directional** – prediction of specific order of group differences on DV or specific relationship between IV and DV

Qualitative IV (sex is the IV here):

Males will have higher test scores than females.

Quantitative IV (time spent studying [as measured in hours]):

The more time spent studying, the greater will be one’s final exam test score.

Or

There will be a positive association between time spent studying and final exam test scores.

Or

There will be a negative association between time spent studying and final exam test scores.

**Non-directional** – prediction general group differences on DV or general association between IV and DV

Qualitative IV (sex is the IV):

There will be a difference in test scores between males and females.

Quantitative IV (time spent studying [as measured in hours]):

The will be a relationship between time spent studying and final exam test scores.

**Null** – prediction of no group differences on DV or no relationship between IV and DV

Qualitative IV (sex is the IV):

There will be no difference in test scores between males and females.

Quantitative IV (time spent studying [as measured in hours]):

The will be no relationship between time spent studying and final exam test scores.

**Common Mistakes**

1. Forcing incorrect comparisons, i.e., incorrectly comparing IV to DV; remember, when the IV and qualitative, one compares the DV across categories of the IV

Example

There will be a difference between sex and reading test scores.

Revised

There will be a difference in reading test scores between females and males.

2. Inventing groups to compare that are not part of the variable provided

Example

IV = Types of peer instruction (think-pair-share vs. peer tutoring)

DV = science scores

Those who receive peer instruction will obtain higher science scores than those who do not receive peer instruction.

Revised

Those who participate in peer tutoring will obtain higher science scores than those who participate in think-pair-share.

3. Artificially creating comparisons for quantitative independent variables

Example

IV = Number of hours of studied before science test

DV = science scores

Those who study many hours will obtain higher science scores than those who study few hours.

or

Those who study will obtain higher science scores than those who do not study.

Revised

The more hours one studies, the higher will be obtained science scores.

4. Ambiguous differences; group differences not specified

Example (supposed to be non-directional)

IV = Type of instruction (cooperative learning or self-paced)

DV = student scores

Type of instruction (cooperative learning or self-paced) affects student scores.

also

Type of instruction (cooperative learning or self-paced) is related to student scores.

How are these hypotheses ambiguous; how are they poorly worded?

Hypothesis is ambiguous about possible group differences.

Also, one should not use affects, relates, association, or similar wording when the IV is quantitative.

Do these hypotheses clearly state which group will perform better or whether differences in scores will exist between groups?

Consider this example which is consistent with the poorly worded hypothesis, yet shows no differences among groups.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Pretest Scores | Treatment | Posttest  Scores |  | Gains? | Group Difference? | Instruction Affect DV? |
| 20 | Cooperative Learning | 50 |  | Yes = 30 | No | Yes |
| 20 | Self-paced | 50 |  | Yes = 30 | No | Yes |

Revised:

Student achievement will differ between cooperative learning and self-paced instruction.

Practice Writing Hypotheses

*Qualitative IV*

IV = Type of reward: reading related vs. non-reading related

DV = student motivation to read

Write directional, non-directional, and null for each (post one type at a time).

With qualitative IV one focuses on group differences or group comparisons, not relationships/associations, and be sure to specify all groups involved. Failure to include all groups involved in the IV can lead to confusion or lead to incorrect assumptions about which groups are compared.

Post your Directional ----

IV = Type of reward: reading related vs. non-reading related (Qualitative)

DV = student motivation to read

Directional = Students given non-reading related rewards will have higher motivation to read than comparable students given reading related rewards.

Someone once asked: how can showing gains not be a difference between groups?

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Groups | Pretest of Motivation | Treatment | Posttest  Motivation | Gains? |  |
| A | 20 | Reading Related | 40 | Yes |  |
| B | 20 | Non-reading Related | 40 | Yes |  |

Motivation scale: 1= low, 50 = high

Post your Non-directional ---

IV = Type of reward: reading related vs. non-reading related (Qualitative)

DV = student motivation to read

Non-directional = There will be **a** difference in motivation to read between those given reading related rewards and those given non-reading related rewards.

Post your Null ---

IV = Type of reward: reading related vs. non-reading related (Qualitative)

DV = student motivation to read

Null = There will be **no** difference in motivation to read between those given reading related rewards and those given non-reading related rewards.

*Quantitative IV*

IV = Number of books read weekly by parents to their children (Quant.)

DV = Motivation to read

Write directional, non-directional, and null for each (post one type at a time).

NEVER use the word **difference**, or **make a comparison**, if the IV is quantitative.

Post your Directional ---

IV = Number of books read weekly by parents to their children (Quant.)

DV = Motivation to read

Directional 1 = The **more** books read weekly by parents to their children, the **greater** the motivation to read

Directional 2 = There is a positive relation between number of books read weekly by parents to their children and children’s motivation to read.

Directional 3 = As the number of books read weekly by parents to their children increases, the children’s motivation to read is expected to decline.

Post your Non-directional ---

IV = Number of books read weekly by parents to their children (Quant.)

DV = Motivation to read

Non-directional = There is **a** relation between number of books read weekly by parents to their children and children’s motivation to read.

Post your Null ---

IV = Number of books read weekly by parents to their children (Quant.)

DV = Motivation to read

Null = There is **no** relation between number of books read weekly by parents to their children and children’s motivation to read.

1. **Central Tendency**

**9. Variability = Dispersion**

**11. Boxplot or Box and Whisker Plot**

See

Chat 1 - Notes 1 Descriptive and Inferential Statistics (posted in discussion forum)

for SPSS example.

**8. Sampling Error vs. Bias**

Take a sample from a population, and sample is randomly selected, therefore any difference between the statistic and the corresponding parameter is known as sampling error.

Mean Age for Students in this Course

Statistic (from sample of 5 students) = M = 37.333

Population (everyone enrolled in course) = µ = 41.25

Sampling Error = M - µ = 37.333 – 41.25 = -3.917

Bias is systematic discrepancy (not due to random fluctuation) between sample statistic and population parameter, and is usually due to poor study design (how samples are selected; approach to selecting sample is systematically flawed).

**10. Frequencies and Percentile Ranks**

Variable is Sex of students in class, and they are as follows:

M, M, F, F, F, F, M, F, M, F, F

**sex**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | Freq. | Percent  or  Relative  Frequency | Valid Percent  or  Relative  Frequency | Cumulative Percent  or  Cumulative  Relative  Frequency |
| Valid | f | 7 | 63.6 | 63.6 | 63.6 |
| m | 4 | 36.4 | 36.4 | 100.0 |
| Total | 11 | 100.0 | 100.0 |  |

|  |
| --- |
| Note:  \*Percent or Valid Percent is sometimes called Relative Frequency  \*Cumulative Percent is sometimes called Cumulative Relative Frequency |

Sex Freq Relative Freq.

Females 7 .6363 (64%)

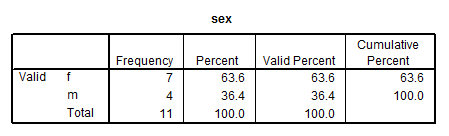
Males 4 .3636 (36%)

N = 11 students,

7/11 = .6363,

4 / 11 = .3636

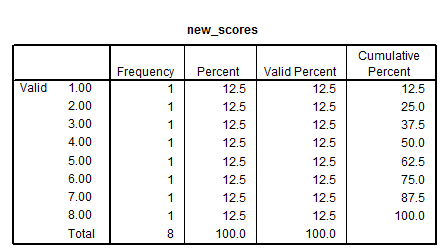
Example of sex with SPSS



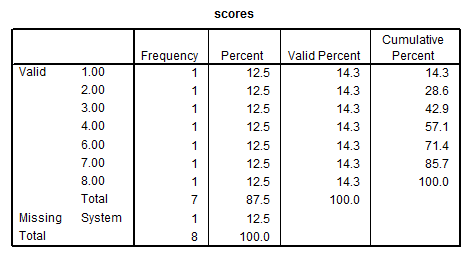
Example: 6, 1, 3, 7, 5, 4, 2, 8

Find frequencies and relative frequencies for the above scores in SPSS

From SPSS



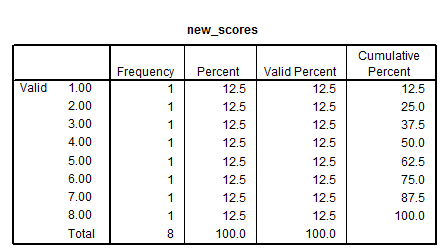
Assume response 5 was omitted from the 8 observations – note difference in percent vs. valid percent.



Percentile Rank

Column cumulative percent = percentile rank for raw data.

What is a percentile rank?

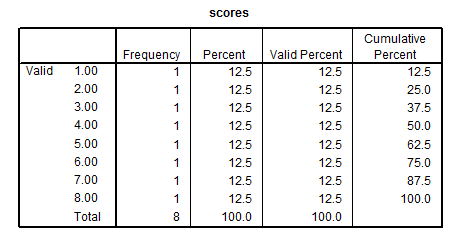


Most common definition and the one we will use:

PR = percentage (or proportion) of scores **at or** below a given score

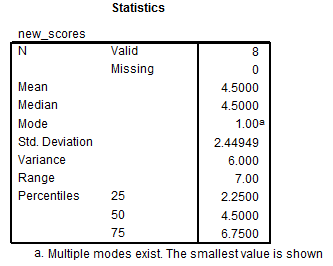
Less common (and we won’t use this one):

PR = proportion (or percentage) of scores below a given score.



Example – for score of 5, 62.5 is the PR which means 62.5% of sample scored 5 or less.

PR = 50 = median



Note that 2.25, 4.50, and 6.75 do not appear in our data. These represent calculated percentile scores for the ranks of 25, 50, and 75. They differ from the values provided by the Cumulative Percent column for the percentiles, and this discrepancy is common for small data files, and the even number of values creates the problem of the median of 4.00 vs. 4.50.

**Quartiles**

Quartiles – what are these?

Quartiles are the 25th, 50th, and 75th percentiles and produce four sections with equal numbers of sampled units in each section.

Divide distribution into 4 sections, with 25% of scores in each section based, upon percentile ranks generally, but specifically using this these formulas:

1st quartile –median between lowest score and overall median of distribution

2nd quartile –median of distribution

3rd quartile –median between highest score and overall median of distribution

Also

1st quartile – 25th percentile

2nd quartile – 50th percentile (median)

3rd quartile – 75% percentile

i.e.,:

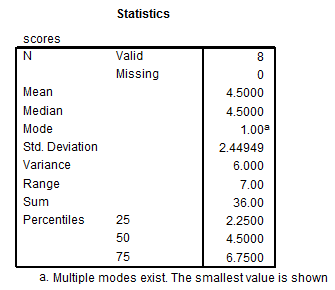
Scores 1 2 3 4 5 6 7 8

↑ ↑ ↑

Quartiles = 1st =2.5 2nd=4.5 3rd=6.5

Percentiles = 25 50 75

SPSS reports different values for quartiles: 2.25, 4.50, and 6.75



There are slight differences in quartile calculation, so if you do it by hand use the formula above and if you rely on software report whatever values they provide because all formulas for quartiles (and percentiles) provide close estimates.

